

Detecting Illicit Drug Trade on SNS through Image Classification and Real-time Monitoring

Tetta Hayashi
Ritsumeikan University
Osaka, Japan

Ryo Nojima
Ritsumeikan University
Osaka, Japan

Abstract—Illegal drug trafficking on social networking services (SNS) has become a growing problem with increasingly sophisticated methods. In Japan, anonymous accounts on platforms such as X (formerly Twitter) frequently post hidden advertisements for illegal drugs among seemingly harmless content. These posts often contain embedded Telegram URLs where the final transaction is concluded. In this study, we develop a prototype of a comprehensive system that detects drug-related content on X, extracts related Telegram links, and visualizes illegal communication networks among users and groups.

I. INTRODUCTION

The evolution of online communication has been paralleled by the rise of cyber-enabled crime. Social networking services (SNS) such as X (formerly Twitter), Instagram, and Telegram offer low-cost, anonymous, and fast channels for information sharing. While these platforms have enabled greater connectivity, they are also exploited for illegal activities, particularly drug distribution. In Japan, drug traffickers frequently use Japanese-language SNS to advertise illicit substances using coded language, obscure pricing schemes, and images with embedded text. These posts often serve as gateways that redirect potential buyers to Telegram groups via URL links.

The current research focuses on detecting this transactional flow from X to Telegram. Traditional moderation systems are largely text-based and ill-suited to the multimodal, multilingual, and evasive nature of drug-related posts. Our approach proposes a hybrid model combining image classification, optical character recognition (OCR), natural language processing (NLP), and hyperlink extraction. The overarching goal is to develop a surveillance system capable of identifying drug-related activity, tracing associated Telegram links, and visualizing the underlying network between users and drug-dealing channels.

Background. In recent years, Japan has witnessed a notable increase in online drug trading, with social media platforms playing a central role in this development. The anonymous and pseudonymous nature of platforms such as X (formerly Twitter) allows users to maintain persistent online identities while evading real-world accountability. A common behavioral pattern observed among individuals involved in illicit drug distribution begins with the posting

of cryptic messages and images that subtly suggest the availability of illegal substances. These posts often include embedded Telegram group links, either clearly presented or disguised within the text or visual elements of the post. Once contact is established, the actual transaction or negotiation typically takes place within encrypted Telegram channels, offering an additional layer of privacy and security.

These cross-platform interactions pose significant challenges for monitoring and detection, as the initial content on X may not directly reference drugs, and the final stages of communication occur on a separate, encrypted platform. Additionally, many of these posts rely on images that contain embedded Japanese text, symbolic emojis, or numeric expressions such as “1/4000,” which implicitly refers to one gram of a substance sold for 4000 yen. Such multimodal and culturally specific encoding complicates the identification of illicit activity through conventional moderation or keyword-based filtering techniques.

Related Work. Many studies have been conducted on the illegal drugs trade on SNS. The authors of [1] proposed a method to identify users engaging in illicit drug transactions on Instagram. The proposed approach uses machine learning to analyze images and text for detection. In [2], to determine whether a user is a drug trafficker on social media, the authors not only treated the content-based features (text and image features), but also the complex relationships among users, posts, and keywords using the graph neural networks. Also in [3], an illicit drug dealer identification method was proposed by constructing a large-scale multimodal dataset. The study of [4] focuses on detecting illegal drug selling posts (not accounts) from hashtag searches and using deep learning to analyze text from posts. The author of [5] provide a systematic study of the fine-grained classification of drug-related hashtags in posts or comments on Instagram. In [6] the authors proposed a method that leverages large language models such as ChatGPT to detect illicit drug trafficking activities on social media. Many of these studies focused on detection on Instagram.

In research related to X and drugs, the paper [7] investigates which types of drugs are prevalent and in which locations within the United States. In [8], to more

accurately understand NUPM (Non-Medical Use of Prescription Medications) behavior on the Internet, the authors monitored and analyzed data on X to clarify the frequency of NUPM-related posts and investigated the actual situation of unauthorized access to drugs through illegal online pharmacies. The authors of [9] conducted experiments using various classification algorithms, including SVM, XGBoost, Random Forest (RF), Naive Bayes (NB), LSTM, and BERT, on a dataset of 156,521 tweets posted between 2008 and 2022 for the purpose of drug use detection.

The promotion of illegal drugs has become increasingly prevalent on popular social media platforms such as TikTok, Instagram, and YouTube. In [10], the authors presented the systematic study on the ecosystem of such cross-platform illicit drug promotion and selling businesses. Especially, the authors' analysis include the cases where the YouTube comment section is used to direct users to locations for drug transactions. Our work has similar structure to [10]. We analyze the cases where X (Twitter) is used to direct users to locations (Telegram) for drug transactions, and it seems to be the most representative form of transaction in Japan.

Our Contributions. We define the problem as a multi-stage illicit communication detection pipeline operating across platforms. The goal is to identify drug-related Japanese posts on X, extract associated Telegram links, and analyze the communication network between X users and Telegram groups. This can be analyzed by decomposing the problem into several subtasks. First, a binary classification is performed to determine whether a given SNS post—based on its image and textual content—is drug-related or benign. Next, hyperlink extraction is conducted to identify Telegram URLs embedded in the post body or within overlaid text retrieved via OCR. Following this, entity resolution is applied to associate multiple posts from the same user and to infer potential membership in Telegram groups. With this information, a network construction step creates a graph where nodes represent users and Telegram groups, and edges indicate discovered link-based associations. Finally, a visualization of the resulting network is produced to reveal structural patterns such as clusters, central hubs, and pathways of information flow. We contribute to the above-mentioned items. The challenge lies in the diversity and ambiguity of posts. Some are explicit in naming drugs or prices; others rely on subtle cues, emojis, or image-only communication. Furthermore, Telegram links can appear in shortened forms, use redirects, or be obfuscated using visual manipulation.

II. HOW TO DETECT ILLICIT DRUG TRADE

A. Framework

We designed and implemented a real-time monitoring system that continuously collects and analyzes posts at

short intervals (every 30 minutes), effectively enabling near real-time detection and visualization of emerging drug-related activities on X (formerly Twitter).

processes them through our classification pipeline, and visualizes the results. The system consists of the modular components as Fig.1

This infrastructure forms the foundation of a practical social media surveillance tool focused on illicit drug trade detection in Japanese-language content.

B. Data Collection and Annotation.

To construct a meaningful and representative dataset, we collected Japanese-language posts from X using Twitter's v2 API. Our primary focus was to capture content related to illicit drug trading, specifically posts that either directly advertise drugs or serve as gateways to Telegram groups where transactions occur. The collection process involved several steps:

Data Collection. We applied multiple search filters based on known indicators of drug-related activity. For the data collection, “手押し” meaning hand-to-hand transaction, “1g”, “1/4000”, and “4500 Yen” and so on have employed as keywords, and also `has:images` and `lang:ja`, `-is:retweet` have employed as filters. By these, we continuously collect the data for two weeks including tweet ID, author ID, timestamp, and geolocation if available. To implement, we developed a Python script that continuously queried the API every 30 minutes using `tweepy`. This interval-based collection allows the system to operate in a quasi real-time mode, providing up-to-date monitoring of potential illicit posts while maintaining API rate compliance. The script stored JSON responses into a structured database.

Annotation. Labeling tweets as drug-related or not requires context sensitivity and domain knowledge. A graduate students the data based on drug slang, pricing formats, embedded Telegram links of the form

`t.me/[a-zA-Z0-9_]{3,}`,

and images depicting pills, powders, packaging, and related items.

To train and evaluate our classifier, we manually annotated a dataset of tweets containing both textual and visual content. The labeling process was conducted entirely by one of the authors to ensure consistency and clarity in the annotation criteria. Each tweet was assigned a binary label—positive (drug-related) or negative (benign)—based on a comprehensive assessment of the tweet's text, image content, and OCR-extracted text. A tweet was labeled as drug-related (label 1) if it explicitly or implicitly indicated illicit drug trade activity. Such posts typically contained keywords or slang (e.g., “手押し”, “1/4000”), numerical expressions representing pricing (e.g., “1g 4500 円”), or embedded Telegram links that suggested further interaction. Visual features such as pills, packaging, or

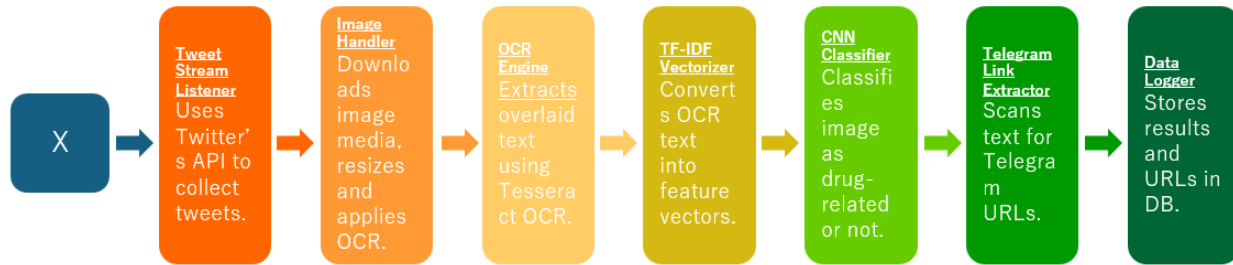


Fig. 1: Framework of Our System



Fig. 2: Examples of images for trading posts on X (in Japanese)

screenshots of drug advertisements also contributed to a positive classification.

C. Detection Method

In Japan, drug transactions are often conducted using images rather than texts in order to evade detection.

To handle the images, we devised a comprehensive preprocessing pipeline to standardize inputs before classification such as resizing images to 700×700 pixels with each pixel normalized to the range $[0, 1]$.

Given the limited number of drug-related posts available for training, we implemented a set of data augmentation techniques to artificially expand the dataset. These techniques included horizontal and vertical flipping, random rotation within a range of ± 15 degrees, adjustments to blur and brightness, as well as cropping and zoom-in transformations. As a result of this augmentation pipeline, the dataset size was increased by approximately a factor of three, yielding around 10,000 images for model training. This enhanced dataset helped improve the robustness and generalization of the classification model, particularly in

recognizing varied visual patterns associated with illicit activity.

We explored two complementary approaches to detect drug-related content on social media: CNN + OCR, and Multimodal Extension (OCR + NLP). In the first method, a Convolutional Neural Network (CNN) was employed to classify images, while Optical Character Recognition (OCR) was applied to extract embedded text for further analysis.

In the second method, we extended the pipeline by incorporating Natural Language Processing (NLP) techniques. Specifically, OCR-extracted text was vectorized using TF-IDF and fused with image features from the CNN model. These enhancements were particularly evident in cases where drug-related cues were present only in the text but not visually obvious.

CNN + OCR. To improve detection accuracy, we employed Tesseract OCR [11], configured for Japanese, to extract overlaid text from tweet images. This process proved essential in uncovering Telegram URLs that were visually embedded within screenshots, stickers, or promotional flyers—elements that are often used to disguise illicit content from simple keyword-based detection methods. We employed Convolutional Neural Network (CNN) architecture optimized for binary classification between drug-related and benign social media images. The model was trained using the Adam optimizer with binary cross-entropy as the loss function. Training was conducted with TensorFlow 2.9 and Keras on an NVIDIA RTX 3060 GPU, and convergence was achieved in approximately 96 minutes.

Multimodal Extension (OCR + NLP). While our primary detection method is based on image classification, we observed that combining visual and textual cues significantly improves detection accuracy. Therefore, we implemented a multimodal extension that integrates image features with text extracted via OCR and processed using TF-IDF.

To enhance detection performance, we applied Tesseract OCR (configured for Japanese) to each tweet image in

order to extract overlaid text. Common examples included Telegram links embedded in screenshots or stickers, numeric expressions such as “1/4000” indicating dosage and price, and stylized slang terms used for promotion. The extracted text was cleaned and normalized to remove noise and irrelevant characters.

Next, the OCR text was converted into sparse feature vectors using the TF-IDF (Term Frequency–Inverse Document Frequency) method. For Japanese text processing, we utilized MeCab with the IPADIC dictionary, excluding stopwords and limiting the vocabulary to the top 5,000 informative terms from the training data.

To perform multimodal classification, we concatenated each OCR-based TF-IDF vector with the image feature vector extracted from the penultimate layer of the CNN model. The combined feature representation was then fed into a dense binary classifier.

D. Experimental Results.

To evaluate the performance of our CNN-based image classification system, we compared it against two common baseline models used in previous research on illicit content detection. The first baseline is a traditional text-based model using TF-IDF feature vectors in combination with a Support Vector Machine (SVM) classifier. The second is a fine-tuned Japanese BERT model applied to tweet text for binary classification.

We employed a training-validation-testing split of 70%–15%–15% to ensure balanced evaluation. The best-performing model, which combined visual and textual features, achieved an average F1-score of 0.964 on the held-out test set, with a recall of 0.972 for the positive (drug-related) class. This multimodal approach—integrating CNN-based image classification with OCR-extracted text vectorized using TF-IDF—led to a 2.1% improvement in F1-score and a 3.4% increase in recall compared to the image-only model. These gains were particularly significant in cases where drug-related cues appeared exclusively in the text and were not visually apparent in the image itself.

To further evaluate the model’s robustness, we assessed performance on datasets of varying sizes using standard metrics such as accuracy, precision, recall, and F1-score. As summarized in Table I, all metrics improved with increased training data, indicating that the CNN classifier benefited from larger datasets.

In addition, our proposed CNN + OCR model outperformed both baselines in F1-score. Text-only models were limited in capturing embedded visual signals like Telegram links or drug imagery, while image-only models struggled when explicit visual cues were absent. By combining both modalities, our method achieved superior overall performance.

However, the OCR component did present challenges, particularly when dealing with noisy, low-resolution, or stylized images. Despite these limitations, even partial text

extraction provided valuable signals, demonstrating the utility of TF-IDF as a lightweight yet effective complement to image-based classification.

Results. All models were evaluated on the same dataset using standard classification metrics: precision, recall, and F1-score. Table II summarizes the comparative results.

Our proposal outperformed all baseline models in terms of F1-score. Notably, text-only models failed to detect key visual cues such as embedded Telegram links or drug-related imagery. While image-only models showed strong performance, they lacked contextual understanding in cases where visual clues were not explicit. By combining image classification with OCR-extracted text, our approach achieved the most robust and comprehensive detection results.

E. Difficulties

Despite the overall strong performance of our model, an error analysis was conducted to uncover its limitations and failure points. In some cases, benign posts were mistakenly classified as drug-related. These false positives typically involved posts about over-the-counter pharmaceutical products, advertisements containing numerical pricing unrelated to drugs (such as event tickets), or unofficial promotions that included ambiguous slang. Such posts often bore visual similarities to actual drug-related content, featuring elements like pill-shaped icons or numeric expressions such as “1/4000.”

Conversely, the model also failed to detect certain drug-related posts. These false negatives often involved blurry or handwritten Telegram links, densely composed meme-style images cluttered with slang or obscure references, and creatively obfuscated content using regional dialects or emojis. These examples highlight the challenges of accurately detecting subtle or noisy cues in real-world imagery and underscore the need for more robust multimodal analysis. These cases highlight the difficulty of detecting subtle or noisy cues in real-world imagery and suggest that incorporating more robust OCR and language models may help reduce such errors. False positives risk mislabeling innocent users, while false negatives result in missed detections of illicit activity. Improving interpretability and adding multimodal signals (text + image) are promising directions for reducing these issues.

- **Tweet Stream Listener:** Uses Twitter’s Streaming API to collect tweets matching predefined filters and keywords in real time.
- **Image Handler:** Downloads image media, resizes to model dimensions, and applies OCR processing.
- **CNN Classifier:** Classifies each image as drug-related or not using the trained CNN model.
- **Telegram Link Extractor:** Scans tweet text and OCR output for Telegram URLs using regex matching.
- **Data Logger:** Stores metadata, classification results, OCR text, and URLs in a structured database.

Dataset Size	Accuracy	Precision	Recall	F1-score
500	0.88	0.85	0.87	0.86
2500	0.93	0.92	0.93	0.92
5000	0.96	0.95	0.96	0.96
10000	0.975	0.97	0.972	0.971

TABLE I: Classification performance on varying dataset sizes

Model	Precision	Recall	F1-score
TF-IDF + SVM	0.82	0.79	0.80
Japanese BERT	0.91	0.87	0.89
Ours (CNN + OCR)	0.97	0.972	0.971

TABLE II: Comparison with baseline models

III. ANALYSIS OF TELEGRAM NETWORK

A. Extraction of Links

One of the key contributions of this study is the detection and analysis of Telegram group links embedded in drug-related posts on X. To extract these links, we applied regular expressions to both tweet text and OCR-extracted image text. The main pattern used was:

(https?://)?(www.)?t.me/[a-zA-Z0-9_]{3,}.

This pattern allowed us to detect fully qualified links (e.g., <https://t.me/abc123>), abbreviated forms (e.g., t.me/groupname), and obfuscated versions embedded in images.

A significant number of Telegram links were embedded visually in image content such as screenshots, stickers, or promotional flyers. We extracted these using Tesseract OCR configured for Japanese and post-processed the output with regex matching.

Extracted links were normalized to remove trailing slashes, case differences, and tracking parameters. Redirect links using services like `bit.ly` and `goo.gl` were resolved using HTTP header inspection.

Out of a total of 10,732 tweets that included images, we identified 2,384 tweets containing at least one valid Telegram link. Among these, 71% of the links were found directly in the tweet text, while the remaining 29% were detected only through OCR applied to the images. These findings emphasize the necessity of multimodal extraction techniques, especially when facing obfuscated or hidden information in illicit drug-related posts.

B. Network Visualization

To analyze the communication patterns between X users and Telegram groups, we constructed a bipartite graph using the extracted Telegram URLs. The graph was implemented with the NetworkX Python library, and equivalent URLs (e.g., `t.me/group1` and `https://t.me/group1/`) were normalized to a single node.

The resulting graph consisted of 434 nodes in X user and 372 nodes in Telegram user, with a total of 372 edges. The average node degree was 1.71, the graph density was

0.0020, and the clustering coefficient was 0.0, indicating the absence of triangle formations.

From this graph, we observed that most Telegram groups were connected to only one or two users. In contrast, a small number of highly active users were linked to multiple groups, forming star-shaped subgraphs. Overall, no clear community structure was found, suggesting a decentralized pattern of interaction.

Figure 3 illustrates a force-directed layout of the graph, where red nodes represent Telegram users and blue nodes represent X posts.

This network representation helps identify influential users, link redundancy, and the scale of cross-platform interaction patterns.

IV. FUTURE WORK

Building on the results and limitations identified in this study, several directions for future research and system improvement are proposed.

While our current system focuses on Japanese-language content, similar illicit drug trade behaviors exist in other linguistic contexts. Extending the pipeline to support multilingual OCR and NLP (e.g., Korean, Chinese, English) would enable broader regional surveillance.

Incorporating time-series analysis may help identify behavioral patterns such as periodic promotions, user migration across platforms, or coordinated campaigns. Combining account history and content evolution could improve prediction and profiling.

With proper safeguards and ethical oversight, the system could be adapted for use in real-world investigations. Automatic flagging of high-risk Telegram groups and coordinated visualization of network evolution could assist investigators in mapping trafficking infrastructure.

Using unsupervised clustering or co-occurrence mining, the system could learn to identify emerging slang or codewords associated with drug trafficking. This would reduce reliance on manually curated keyword lists and adapt to linguistic drift.

Future versions of the system could monitor multiple platforms (e.g., Instagram, LINE, Discord) and fuse results into a unified graph. This would allow tracing the full path of illicit communications from public posting to private transaction.

V. CONCLUSION

In this study, we proposed and implemented a multimodal detection framework designed to identify and

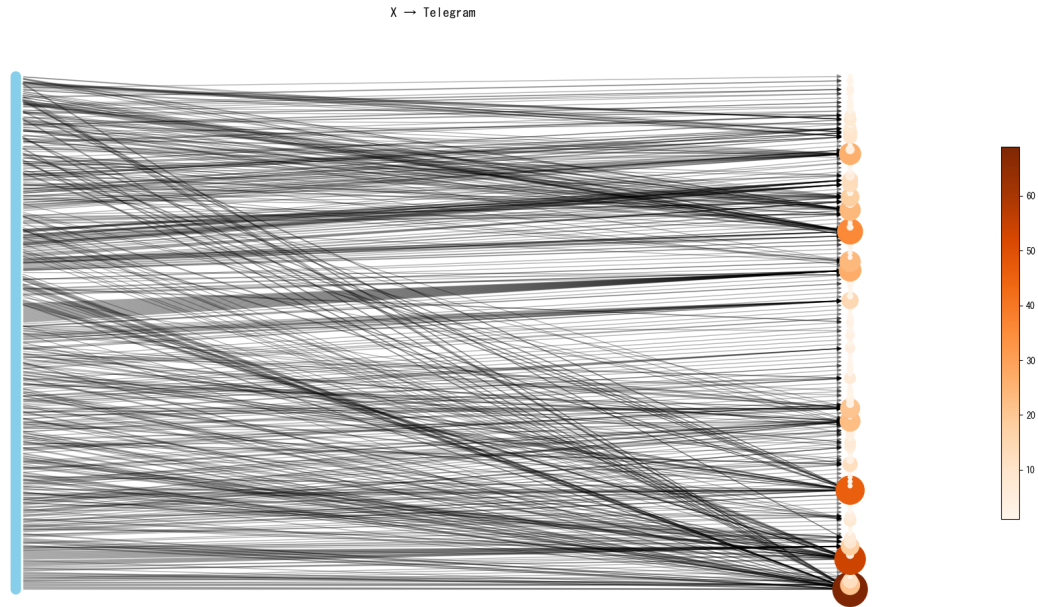


Fig. 3: Bipartite network of X users and Telegram group links

analyze illicit drug trade activities occurring on Japanese-language social media platforms, with a particular emphasis on the interaction between X (formerly Twitter) and Telegram. Our work addresses a growing societal concern: the migration of criminal activity to decentralized, encrypted, and cross-platform digital environments.

The system integrates multiple technologies to support robust, scalable First, we developed a CNN-based image classifier to detect drug-related promotional materials with high precision. Second, we applied Japanese-configured OCR to extract textual content embedded in images, which was further vectorized using TF-IDF to capture linguistic cues such as pricing formats, slang, and contact links. By fusing visual and textual features, the proposed multimodal model achieved superior performance compared to unimodal baselines.

Furthermore, we constructed a bipartite network representing the relationships between X users and Telegram groups, based on the Telegram links extracted and normalized from tweet text and images. Graph analysis revealed structural patterns that reflect the behavior of illicit trade networks, such as decentralized connectivity and the presence of hub-like users. However, in this network, the appearance of a Telegram link in a tweet is implicitly assumed to indicate a “connection” to that group. In practice, such links may not always imply actual membership or engagement, but could instead represent indirect references—such as mentions, criticisms, or public exposure of the group. Therefore, we emphasize that the constructed network does not directly capture confirmed

group membership or active participation, but rather reflects patterns of mention and exposure. Clarifying this distinction enhances the validity of our network-based interpretations and strengthens the overall credibility of the study.

Despite the system’s effectiveness, several challenges remain, including OCR reliability on distorted images, generalization to evolving slang, and integration across additional platforms. We also acknowledge the need for careful oversight in applying such technologies to real-world settings, especially in ways that may influence law enforcement or platform policy.

In conclusion, our work demonstrates a practical and extensible approach to detecting online drug trafficking within the Japanese linguistic and cultural context. It represents a critical step toward automated, ethically responsible cybercrime surveillance tools that can adapt to the shifting landscape of illicit behavior on social media. Future work will aim to scale the system to multiple languages and platforms while improving its resilience to adversarial evasion tactics.

A. Security, Privacy, and Ethical Considerations

This study aims to automatically analyze publicly available social media content for the purpose of detecting illicit drug trade activities. While the tweets and images analyzed in this research are publicly accessible, we recognize the importance of addressing privacy and ethical concerns with due diligence.

Our research is grounded in the principles of cybersecurity research ethics. Only public content from X

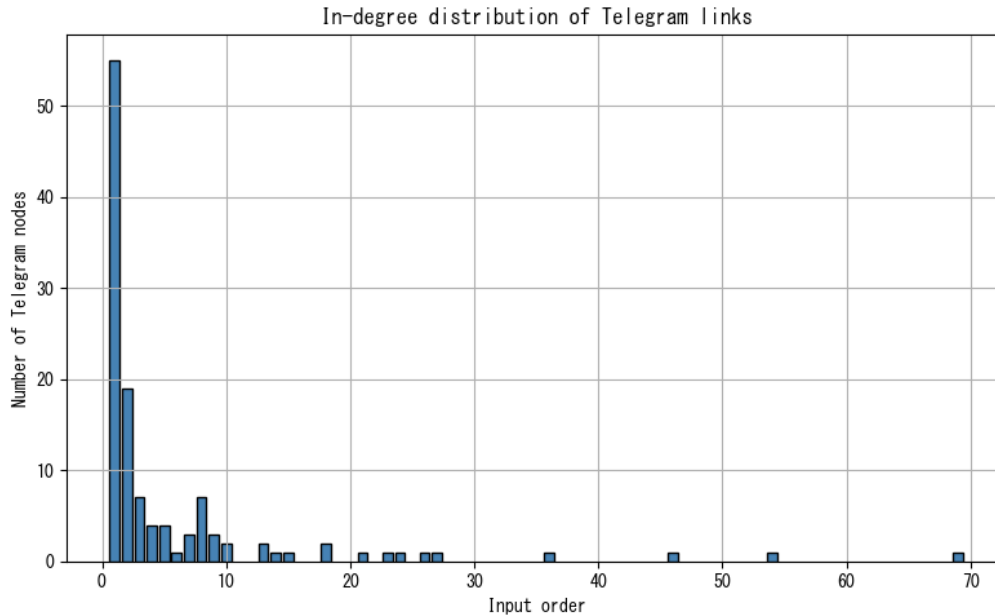


Fig. 4: Degree distribution of Telegram nodes

(formerly Twitter) was collected. No access was made to protected accounts, private accounts, direct messages. Similarly, Telegram link analysis was restricted to links found in publicly posted tweet text or embedded within tweet images. No content from private Telegram groups was accessed or analyzed.

Secondly, the outputs of this system are intended solely to support academic investigation and research-driven monitoring. The system is not designed to trigger punitive actions or surveillance without human oversight. Ambiguous or borderline cases cannot be automatically acted upon and require manual verification by human analysts. By adhering to these safeguards, our research maintains ethical integrity while addressing a critical social issue.

REFERENCES

- [1] X. Yang and J. Luo, "Tracking illicit drug dealing and abuse on instagram using multimodal analysis," *ACM Trans. Intell. Syst. Technol.*, vol. 8, no. 4, Feb. 2017. [Online]. Available: <https://doi.org/10.1145/3011871>
- [2] Y. Qian, Y. Zhang, Y. F. Ye, and C. Zhang, "Distilling meta knowledge on heterogeneous graph for illicit drug trafficker detection on social media," in *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds., vol. 34. Curran Associates, Inc., 2021, pp. 26 911–26 923. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2021/file/e234e195f3789f05483378c397db1cb5-Paper.pdf
- [3] C. Hu, M. Yin, B. Liu, X. Li, and Y. Ye, "Identifying illicit drug dealers on instagram with large-scale multimodal data fusion," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 5, Sep. 2021. [Online]. Available: <https://doi.org/10.1145/3472713>
- [4] J. Li, Q. Xu, N. Shah, and T. K. Mackey, "A machine learning approach for the detection and characterization of illicit drug dealers on instagram: Model evaluation study," *J Med Internet*

- Res*, vol. 21, no. 6, p. e13803, Jun 2019. [Online]. Available: <http://www.jmir.org/2019/6/e13803/>
- [5] C. Hu, B. Liu, Y. Ye, and X. Li, "Fine-grained classification of drug trafficking based on instagram hashtags," *Decision Support Systems*, vol. 165, p. 113896, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167923622001671>
- [6] C. Hu, B. Liu, X. Li, Y. Ye, and M. Yin, "Knowledge-prompted chatgpt: Enhancing drug trafficking detection on social media," *Information & Management*, vol. 61, no. 6, p. 104010, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0378720624000922>
- [7] C. Buntain and J. Golbeck, "This is your twitter on drugs: Any questions?" in *Proceedings of the 24th International Conference on World Wide Web*, ser. WWW '15 Companion. New York, NY, USA: Association for Computing Machinery, 2015, p. 777â 782. [Online]. Available: <https://doi.org/10.1145/2740908.2742469>
- [8] T. Katsuki, T. K. Mackey, and R. Cuomo, "Establishing a link between prescription drug abuse and illicit online pharmacies: Analysis of twitter data," *J Med Internet Res*, vol. 17, no. 12, p. e280, Dec 2015. [Online]. Available: <http://www.jmir.org/2015/12/e280/>
- [9] R. G. AlGhannam, M. Ykhlef, and H. Al-Dossari, "Leveraging ensemble method with transformer for robust drug use detection on twitter," *Preprints*, July 2023. [Online]. Available: <https://doi.org/10.20944/preprints202307.1577.v1>
- [10] M. Zha, Z. Lin, S. Tang, X. Liao, Y. Nan, and X. Wang, "Understanding cross-platform referral traffic for illicit drug promotion," in *Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 2132â 2146. [Online]. Available: <https://doi.org/10.1145/3658644.3670383>
- [11] R. Smith, "An overview of the tesseract ocr engine," in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, vol. 2, 2007, pp. 629–633.